N8103-149/150/151/160
RAID Controller

N8103-156
MegaRAID® CacheCade®

Feature Overview

April 2012
Rev.1.0

NEC Corporation

Contents

1 Introduction 3

2 Types of RAID Controllers 3

3 New Features of RAID Controllers 5

3.1 Improvements in the basic functions of RAID controllers 5

3.2 HDD/SSD support 6

3.3 RAID 60 support 6

3.4 HDD Power Management Function 7

3.5 Compatibility with EXPRESSSCOPE Engine 3 9

4 MegaRAID CacheCade 11

4.1 What is MegaRAID Cache? 11

4.2 Implementation 12

4.3 Performance 13

4.4 Hot spot data 14

5 Conclusion 18

1 Introduction

N8103-149/150/151/160 RAID controllers, which support NEC Express5800 Server Series, employ DDR3 Memory for cache memory and a dual core IO processor from LSI. Compared to existing RAID controllers with RAID Access functions, sequential reading performance is doubled and sequential writing is improved 1.6 times. (Tested by NEC)

MegaRAID® CacheCade® enhances RAID performance by allocating a Solid State Drive (SSD) as cache memory for the RAID controller.

This paper is an overview of new functions for the RAID controller, and also covers in detail the MegaRAID CacheCade.

2 Types of RAID Controllers

In the Express5800/100 series, the RAID controllers in Figure 1 below are products for sale starting in 2012.

N8103-149 is the most suitable RAID controller for customers building a small scale RAID configuration of RAID 0, 1, and 10.

N8103-150 and N8103-151 are RAID controllers that support all general RAID levels which use RAID 0, 1, 5, 6, 10, 50, and 60.

N8103-160 is a RAID controller for connecting external storage, and supports RAID 0, 1, 5, 6, 10, 50, and 60.

Figure 1. Specifications of RAID controller

Model Name	Product Name	Type	RAID Level	Cache Capacity	Battery	MegaRAID CacheCade
N8103-149	RAID Controller (512MB, RAID 0/1)	For built-in use (8 Ports)	RAID 0, 1, 10 (compatible with SATA/SAS)	512MB	Supported	Unsupported
N8103-150	RAID Controller (512MB, RAID 0/1/5/6)		RAID 0,1,5,6 10,50,60 (compatible with SATA/SAS)			
N8103-151	RAID Controller (1GB, RAID 0/1/5/6)			1GB		Supported
N8103-160	RAID Controller (1GB, RAID 0/1/5/6)	For external use (8 Ports)				Unsupported

N8103-149/150/151

N8103-160

3 New Features of RAID Controllers

The RAID Controllers have the following new features:

- Employ a dual core IO processor as the RAID computing processor (equipped with a LSISAS2208 chip).
- Employ DDR3 memory as a cache memory. Equipped with a cache capacity of 512MB or 1GB (existing products have 256MB or 512MB cache).
- Complied with MegaRAID CacheCade, which improves RAID performance by allocating the SSD as RAID cache memory (refer to section 4).
- Support SAS SSD and SATA Hard Disk Drive (HDD), which enables a transfer rate of 6Gb/s.
- Support RAID 60. (Also supports the RAID system management utility Universal RAID Utility).
- Reduce power consumption by employing the HDD power management function (Manage Powersave), which controls the power supply to the spare disk.
- Gather information about the HDD and operating RAID controller, by EXPRESSSCOPE ® Engine 3 (a function of NEC Server Management).

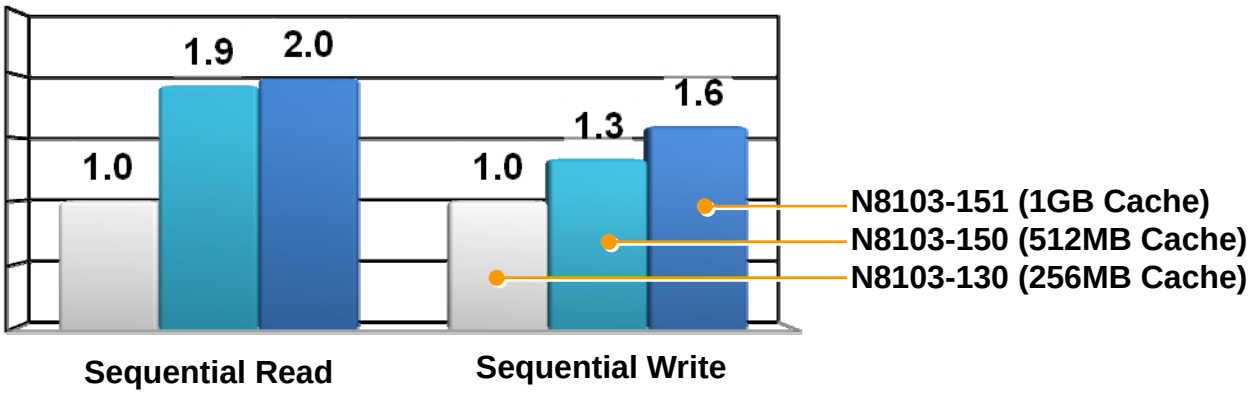
3.1 Improvements in the basic functions of RAID controllers

The RAID controllers employ an LSISAS2208 chip from LSI as the RAID computing processor, and are equipped with 512MB or 1GB cache capacity (existing products [N8103-129/130/134/135] have 256MB or 512MB cache).

In NEC tests using three HDD in a RAID 5 system, the RAID access doubled for sequential reading performance, and increased sequential writing performance by 1.6 times.

Figure 2. Comparing RAID Controller Performance

(Improvement in RAID Access. Performance of the N8103-130 RAID Controller = 1.0)



Measuring Environment (NEC Tests)

Iometer, 3x SAS HDD (10,000rpm), RAID 5, Data Size 8KB, Windows Server 2008

3.2 HDD/SSD support

The RAID Controller now supports SAS SSD and SATA HDD, with a maximum data transfer rate of 6Gb/s.

By supporting media with variation, RAID Controllers provide a solution which can best suit the cluster topology of a client.

Figure 3. List of Supported HDD/SSD.

Supported Drives	RAID Controller (N8103-149/150/151/160)	Existing RAID Controllers (N8103-129/130/134/135)
SATA HDD	***6Gb/s***	3Gb/s
SAS HDD	6Gb/s	6Gb/s
SATA SSD	3Gb/s	3Gb/s
SAS SSD	***6Gb/s***	Unsupported

3.3 RAID 60 support

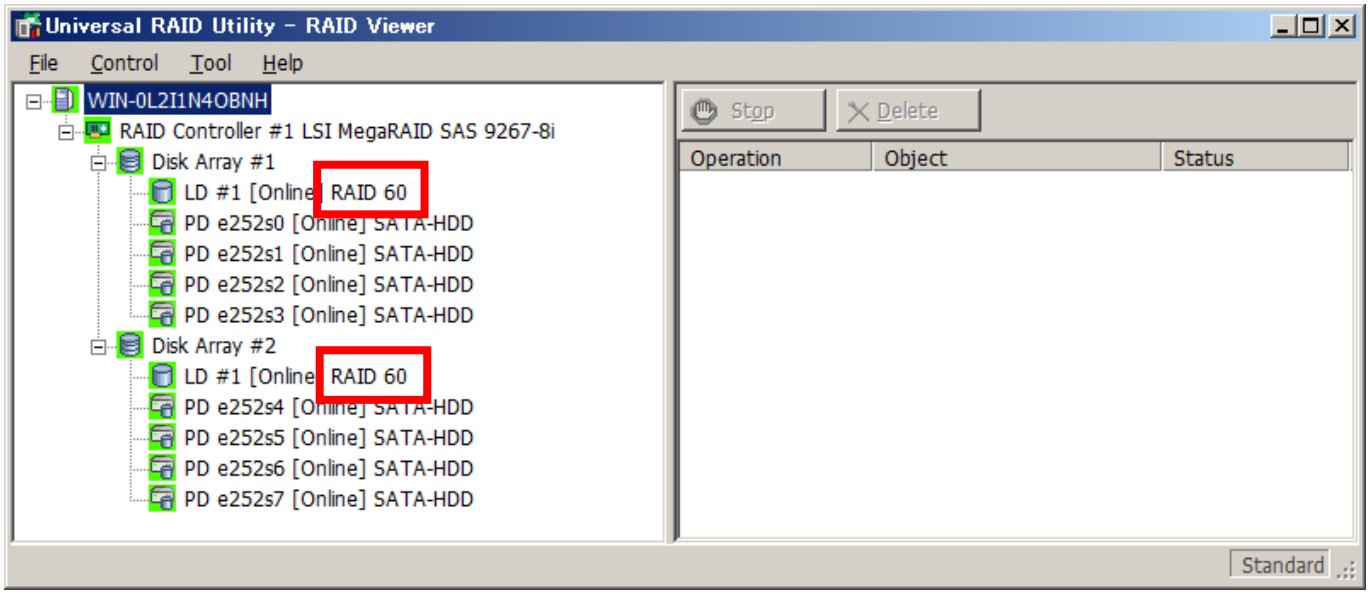
RAID Controllers now support RAID 60.

Monitoring of the RAID controller and HDD is done by using the RAID system management utility “Universal RAID Utility.” It is bundled with a server.

The Universal RAID Utility displays in Japanese when using Windows Server OS, and the interface is compatible with Windows. The RAID system records emergencies or events, which makes it potentially possible to examine and resolve system failures.

The command line interface (CLI) works the same on Linux.

Figure 4. RAID Viewer Screen of Universal RAID Utility.



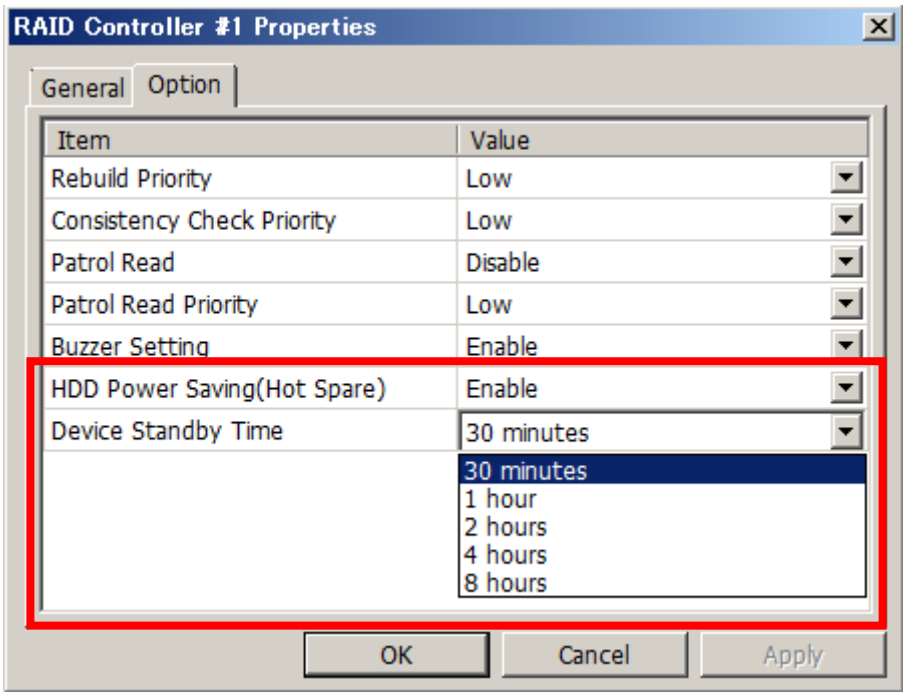
3.4 HDD Power Management Function

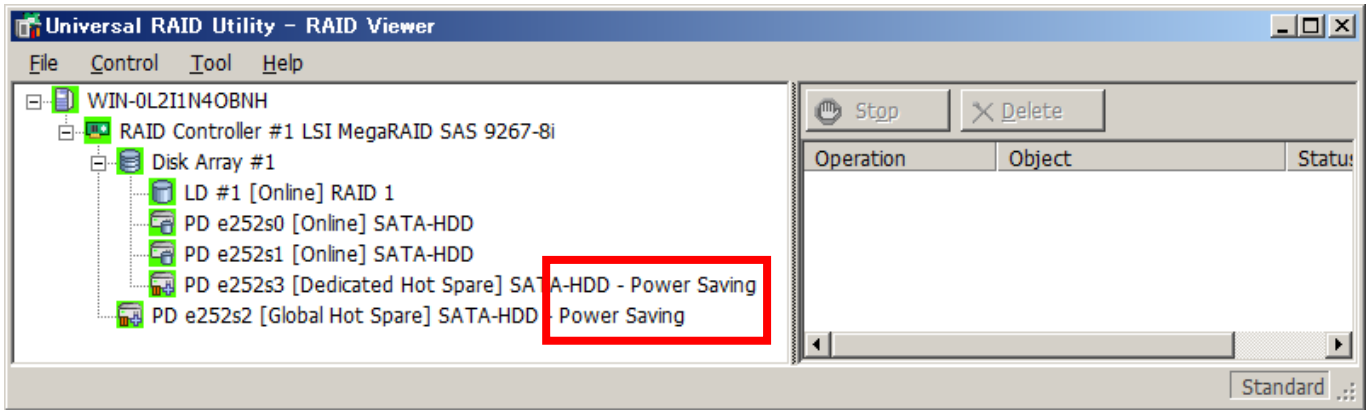
The RAID controller employs a HDD power management function (Manage Powersave), which enables a spare disk to run in a power save mode when the system is operating normally. This reduces power consumption. In general, a spare disk is only required in case of failures and if the system is operating normally; it is not necessary to have it operating continuously.

SATA HDD and SAS HDD consume 5-10W when in standby mode. If 3 spare disks are in the system, 15W could be saved by utilizing the power management function.
As the spare disk is checked periodically by patrol read (a function to check for errors on the entire hard disk drive), errors on the hard disk can be detected early.

The HDD power management function is disabled when shipped. The HDD power management control function must be set to ‘enabled’ from the RAID system management utility “Universal RAID Utility” or the server management utility “ESMPRO/Server Manager." It is possible to configure the transition time to power saving mode. Use the CLI the same way on Linux.

Figure 5. Settings for the HDD Power Management Function (top) and RAID Viewer Screen (bottom)





3.5 Compatibility with EXPRESSSCOPE Engine 3

In addition to the HDD/SSD connection/rebuild/failure monitoring function, the RAID Controllers have a new function that obtains information about the HDD/SSD and RAID Controller when rebooted; independently of the OS (refer to Figure 6 and Figure 7).

The EXPRESSSCOPE Engine 3, a server management function of the NEC Express5800 series, can manage/monitor the server remotely; independen of the OS. As the installation of specific software is not necessary, it eases the management of servers in a remote location (such as in server hosting services).

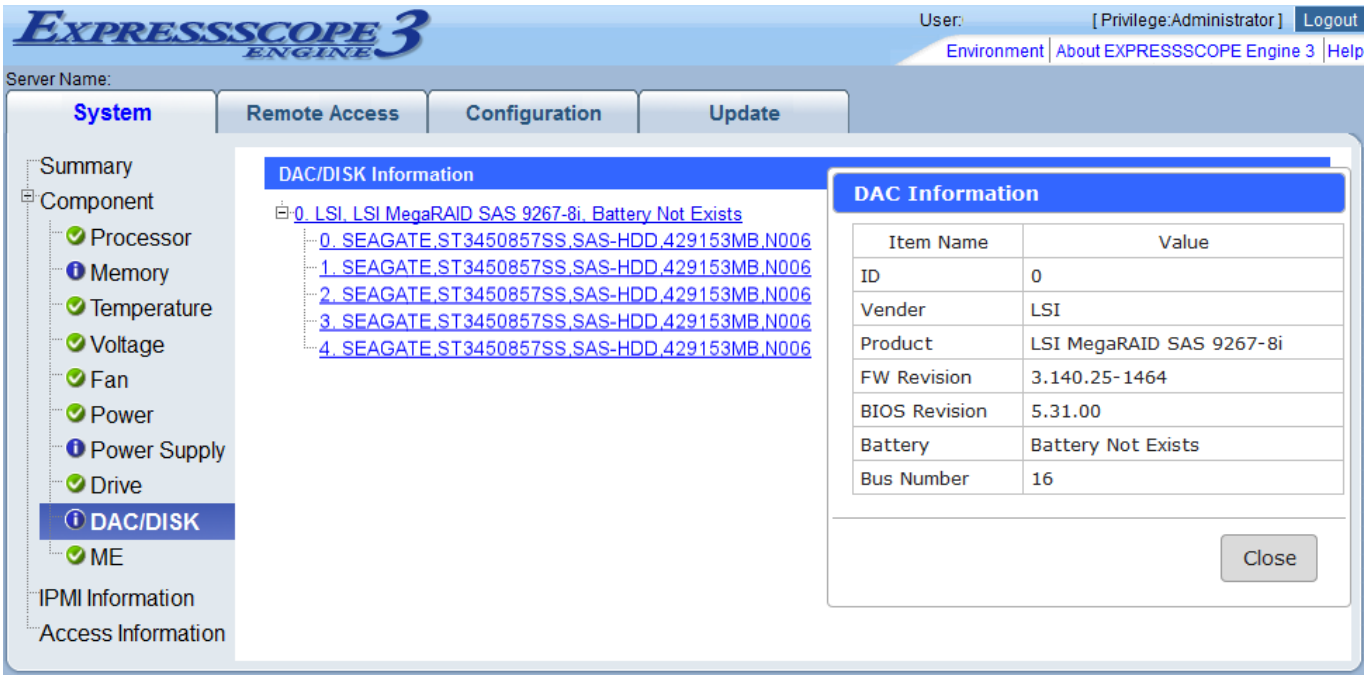
As the HDD/SSD model number and firmware version can be managed remotely, it can also be utilized to manage resources for customers/data center operators who store/manage data in-house.

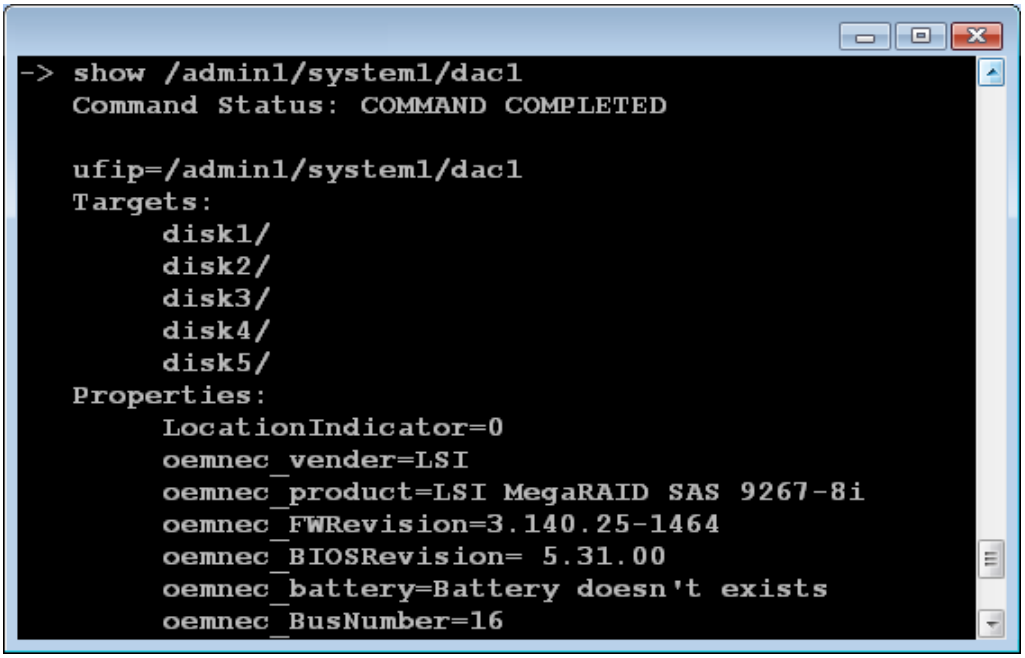
The function for collecting information about the HDD/SSD on EXPRESSSCOPE Engine 3 collects the information upon startup of the server. If the HDD/SSD have expansions or are replaced after startup, the information can be checked after a server reboot.
Be aware that these functions are not predictive monitoring of failures.

Figure 6. Compatible Functions of EXPRESSSCOPE Engine 3 (obtainable independently of OS)

HDD/SSD Failure Monitoring	**HDD/SSD Connection, Rebuilding, Failure Monitoring**
Obtaining Information about the HDD/SSD (New Function)	**Types of HDD/SSD, Vendor Name, Model No., Capacity, Firmware Version**
Obtaining Information about RAID Controller (New Function)	**Vendor Name of RAID Controller, Card Name, Firmware/BIOS Version, Battery Connection, Bus No.**

Figure 7. RAID Controller for EXPRESSSCOPE Engine 3, HDD Information Display (Top: browser, bottom: ssh)





4 MegaRAID CacheCade

4.1 What is MegaRAID Cache?

MegaRAID CacheCade is a function which improves the performance of the random read process, using the SSD as a read cache.

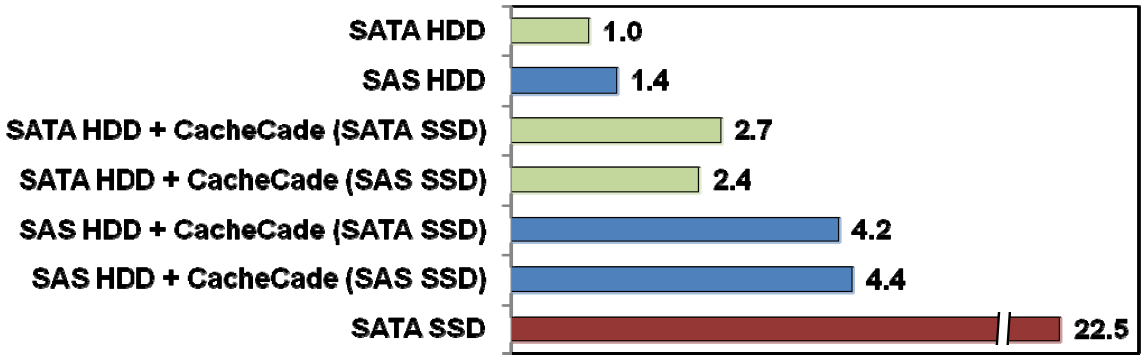
Normally, SSD performs better than HDD, but the cost of implementation per device is high. Cost/performance may be prohibitive when building a RAID system with SSD.

When the MegaRAID CacheCade uses the SSD as a cache for the RAID controller, the RAID system on the HDD can improve performance while keeping the cost to a minimum.

NEC performance analysis results show performance improved up to 3 times when compared with a RAID system with only a HDD (Refer to Figure 8).

Figure 8. Functional Improvements when using MegaRAID CacheCade

(Functional Improvements - performance without CacheCade (when using SATA HDD = 1.0)



Measuring Environment (NEC tests)

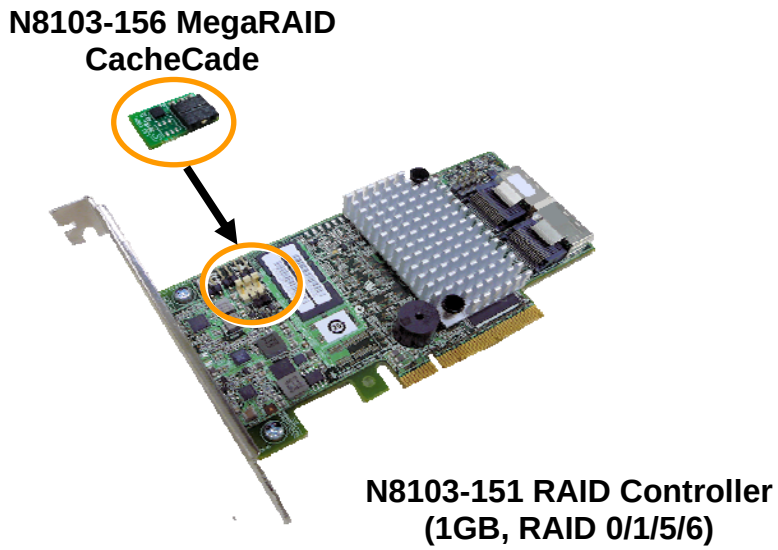
- 2x Intel Xeon E5520, 2GB Memory, equipped with N8103-151and N8103-156
- SATA HDD System: 6x 1TB (7.2K) SATA HDD (RAID5 Configuration)
- SAS HDD System: 6x 600GB (10K) SAS HDD (RAID 5 Configuration)
- The cache for CacheCade uses 100GB SATA or SAS SSD
- SATA SSD System: 6x 100GB SATA SSD
- Settings to collect hot spot data within 90GB
- Expected Usage of Web File Server (I/O Size: 8KB) [4.3Reference]
- OS used is Windows Server 2008

4.2 Implementation

MegaRAID CacheCade is enabled by combining the N8103-156 MegaRAID CacheCade with the N8103-151 RAID controller (1GB, RAID 0/1/5/6). N8103-156 comes with a hardware key that enables MegaRAID CacheCade. Use the key to enable MegaRAID CacheCade and implement to N8103-151.

A BTO shipment includes a hardware key.

It is possible to use a MegaRAID CacheCade, regardless of whether the OS is Windows or Linux, or using VMware.



With MegaRAID CacheCade, a maximum of 512GB can be allocated as cache. When using a 100GB SATA SSD, a maximum of 6 devices can be allocated as caches. SSD supports both SATA and SAS, but they cannot be used together. MegaRAID CacheCade is not recognized as a disk from the OS.

After connecting the SATA or SAS SSD, use the Universal RAID Utility in RAID management utility or Web BIOS to allocate as MegaRAID CacheCade.

MegaRAID CacheCade is primarily a function which improves the performance of random reads. In RAID systems with a lot of sequential read and writes, no such improvement in performance is expected. In order to enhance the random read performance, it is important to store data efficiently on the SSD used for the cache (refer to 4.4).

4.3 Performance

Microsoft defines server applications and workloads. Refer to Figure 9 below.

In the workload categories, the following trends are the results of NECs performance analysis on the MegaRAID CacheCade:

- Improvement in the performance of applications centered on random read, such as Web servers.
- Reduction in the performance of applications with write or sequential reading.

The improvement from MegaRAID CachCade varies depending on the environment and other conditions. NEC recommends analyzing the information in Figure 9 below when implementing.

Figure 9. Performance Improvements for RAID access, via MegaRAID CacheCade per application.

(Performance improvement rate. Performance without CacheCade = 1.0)

http://msdn.microsoft.com/en-us/windows/hardware/gg487522 Virtual Hard Disk Performance

Workload Category	I/O Size	Workload configuration		Rate of improvement in functions when using CacheCade	
				SAS HDD	SATA HDD
Web File Server	4KB	95% Read 5% Write	75% Random 25% Sequential	***1.29***	***1.24***
	8KB			***3.09***	***2.67***
	64KB			***2.06***	***2.16***
Web Server Log	8KB	Write	Sequential	0.93	0.89
OLTP (Online Transaction Processing) DB	8KB	70% Read 30% Write	Random	***1.04***	***1.13***
Decision Support System DB	1MB	Read	Random	0.47	0.90
SQL Server Log	64KB	Write	Sequential	1.03	0.63
Exchange Server	4KB	67% Read 33% Write	Random	***1.07***	***1.06***
Video on Demand	512KB	Read	Random	0.61	0.94
Workstation	8KB	80% Read 20% Write	80% Random 20% Sequential	***1.19***	***1.25***
OS Paging	64KB	90% Read 10% Write	Sequential	0.70	0.76

Measuring Environment (NEC Tests)

- 2x Intel Xeon E5520, 2GB Memory, equipped with N8103-151and N8103-156
- SAS HDD System: 6x 600GB (10K) SAS HDD (RAID 5 Configuration)
- SATA HDD System and 6x 1TB (7.2K) SATA HDD (RAID 5 Configuration)
- The cache for CacheCade is suitable for one 100GB SATA SSD device
- Settings to collect hot spot data within 90GB
- OS used is Windows Server 2008

4.4 Hot spot data

As the MegaRAID CacheCade system improves performance by allocating sections frequently accessed in the HDD (hot spot data) to SSD cache, it is important to analyze to what extent hot spot data is being used.

Steps to analyze hot spot data on Windows Server 2008:

1. Download the Windows Performance Toolkit

   Use to analyze the hot spot data. Windows 7 SDK: download the ISO (GRMSDK_EN_DVD.iso (567.3MB)) from the following URL.
   http://www.microsoft.com/download/en/details.aspx?displaylang=en&id=8442

2. Install the Windows Performance Toolkit

   ISO (burn the CD or use ISO image mounting tools) to install from SDK:
   \Setup\WinSDKPerformancetoolkit\wpt_x86.msi

3. Collect the performance data
   Start from a command prompt and follow the steps below.

   Start collecting performance data
   When the command below executes, the collection of performance data starts. Make sure that the business application for analysis is running before executing the command.

   xperf -on DISK_IO

   Stop collecting performance data
   When the command below executes, the performance data is stored in the specified file and the data collection stops.

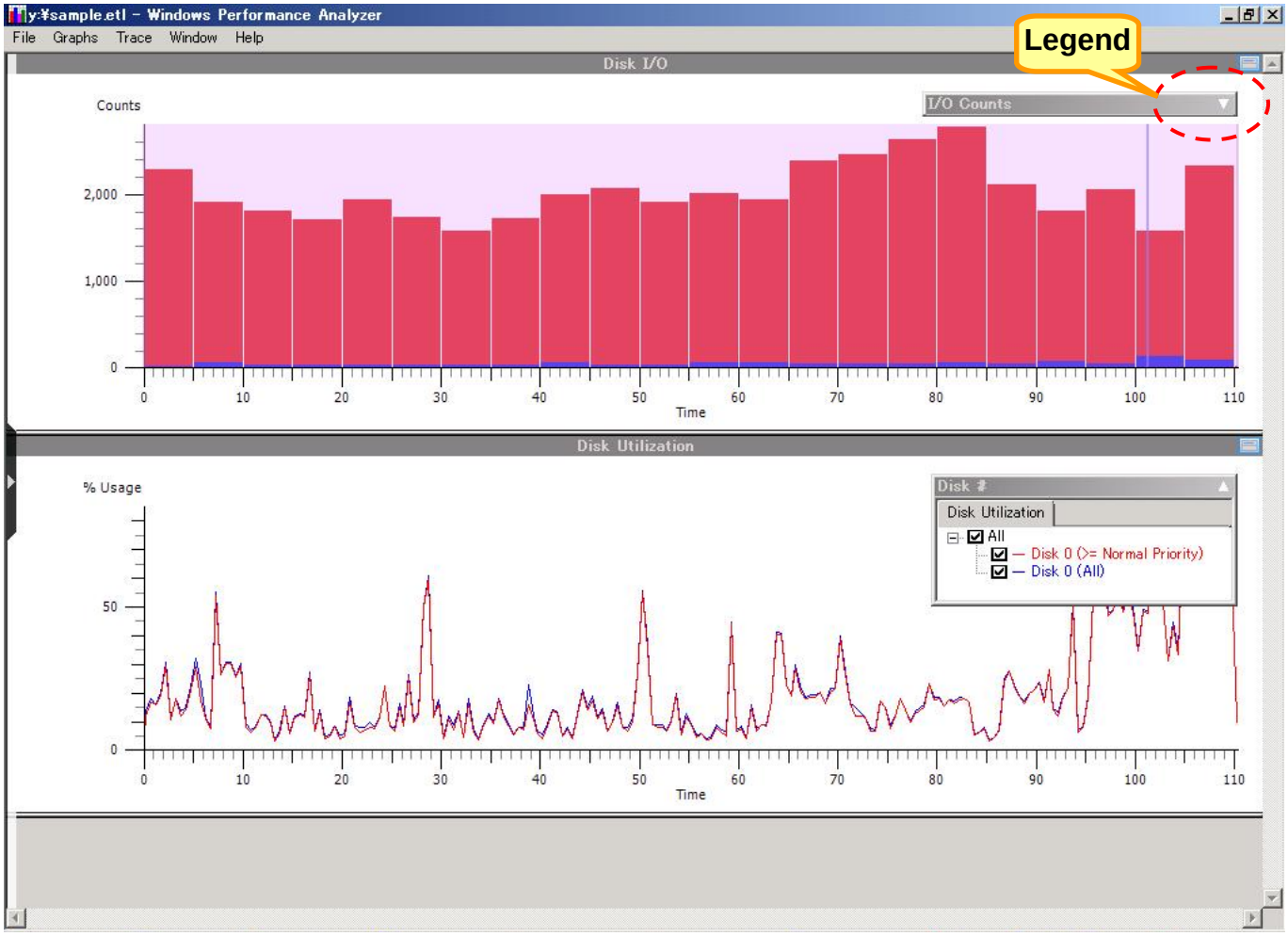
   xperf -d filename (specify the file name for storing the performance data with the full path)

4. Graph and analyze the performance data
   When the command below executes, the specified performance data file is analyzed and displays in a graph.

   xperf filename (specify the file in which the performance data has been stored in Step 2)

   It may take time to display the graph depending on the amount of data. For first time, the following graph (Figure 10) displays.

Figure 10. Conditions for accessing the I/O via Windows Performance Analyzer.



In the above graph, the distribution of read/write access display. The disk usage displays in the graph below. The legend for each graph displays by clicking on the pull down menu in the top right of each graph.

Next, the context menu (Figure 11) displays by right clicking the graph, and by selecting Detail Graph. The time vs. access position on the graph displays in physical disk units (Figure 12).

Select the desired disk from the pull-down menu on the left, and confirm the distribution of access.

Figure 11. Windows Performance Analyzer Context Menu

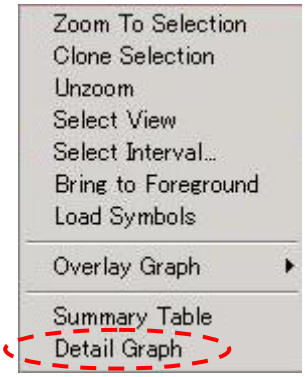
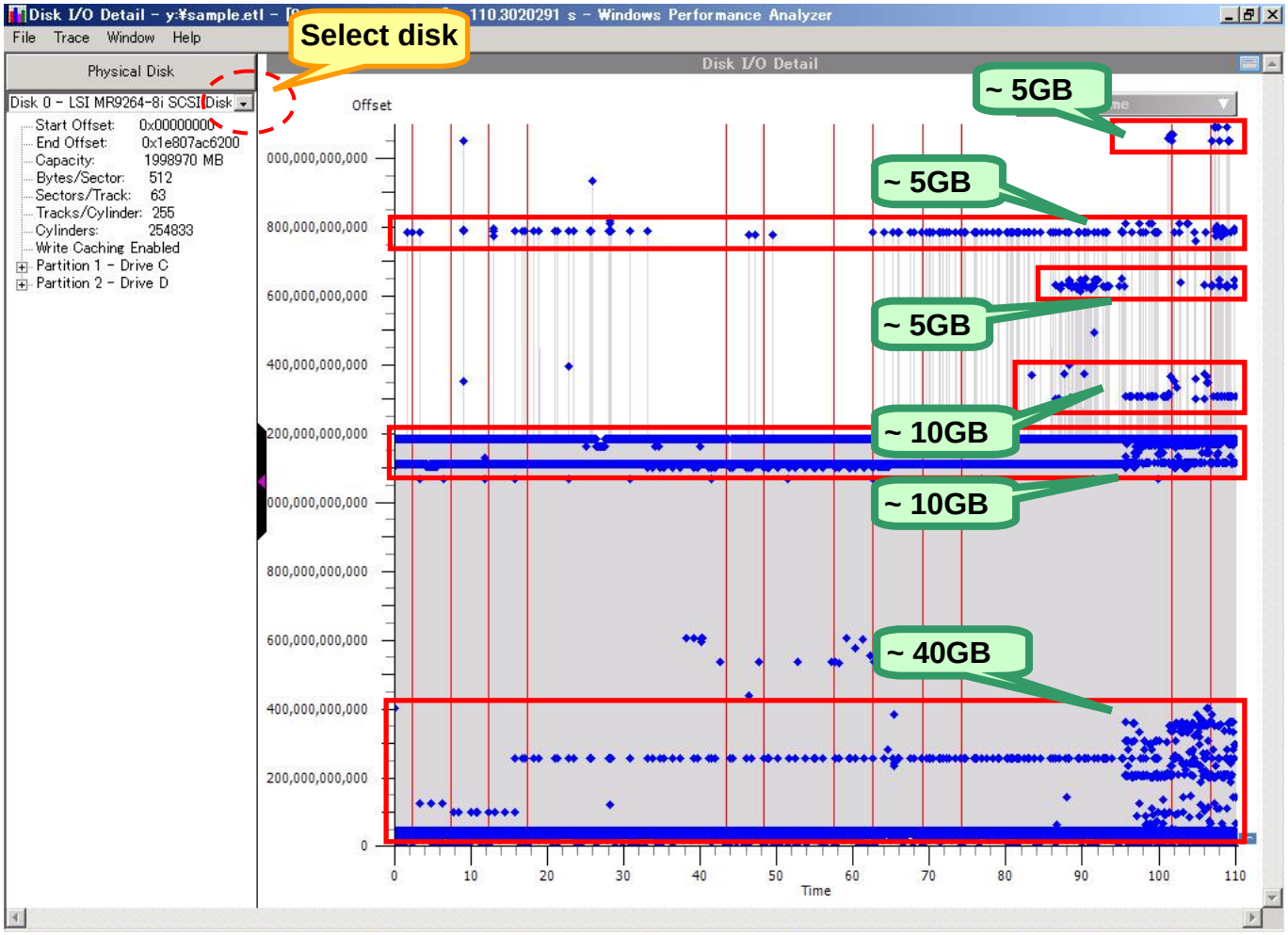


Figure 12. Hot Spot Data



In Figure 12, the horizontal axis shows the time, the vertical axis shows the capacity of the logical drive, and the blue points show the positions of access. The part where the blue points are concentrated is the hot spot data.

In Figure 12, it is possible to estimate the size of the hot spot data as ~ 75GB.

Estimate the size of hot spot data of other logical drives and calculate the total size of the hot spot data. If the size is 150GB, arrange two 100GB SATA SSD devices as data cache for the Mega RAID CacheCade.

5 Conclusion

The usability of RAID Controllers (N8103-149/150/151/160) is improved over existing RAID Controllers due to improved performance and operation management functions.

The RAID Controllers are suitable for the following:

A new cluster topology such as large capacity storage system architecture in a big data environment

A remote operating management essential for data center operations in a cloud computing environment.

Figure 13. Enhancements to RAID Controllers

New Functions	• **Employs a dual core IO processor, and 512MB or 1GB of cache memory for DDR3 to almost double the performance of the RAID system.** • **Improves performance of MegaRAID Cache by allocating the SSD (Solid State Drive) as RAID cache memory** • **Supports SSD and SATA/SAS HDD at 6Gb/s** • **Supports RAID 60** • **Enables management of RAID 60 and CacheCade with the RAID system management utility, Universal RAID Utility** • **Reduces system power by employing the HDD power management control function (Manage Powersave) to control the power supply to the spare disk.** • **Supports the OS-independent (agentless) RAID controller/HDD information function which works in together with EXPRESSSCOPE Engine ® 3, one of the NEC Server Management functions.**

MegaRAID CacheCade speeds up the performance of a HDD array at low cost compared to the RAID configuration using only SSD. However, the degree of improvement can vary greatly depending on applications and the topology of a cluster. Some planning is necessary before introducing.

NEC offers a RAID controller, MegaRAID CacheCade, and SATA SSD testing kit. For more information on obtaining:

Figure 14. Implementing MegaRAID CacheCade.

Advantage	• **Speeds up performance of a RAID system with low cost**
Applications	• **Increased efficiency in applications centered on random read, such as OLTP servers, and web servers for browsing can be expected.** • **Not suitable for applications centered on writing or sequential reading, such as servers collecting Web logs, or streaming services.**
Note:	• **In order to correctly set the capacity for the cache allocated (SSD capacity), it is important to understand the hot spot data (data which is frequently accessed)**

Version History

Rev.1.0	April 2012	First Version Created